

11-13-00

A

**UTILITY PATENT APPLICATION TRANSMITTAL
(Large Entity)**

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Docket No.
ST9-99-180

Total Pages in this Submission

TO THE ASSISTANT COMMISSIONER FOR PATENTSBox Patent Application
Washington, D.C. 20231

Transmitted herewith for filing under 35 U.S.C. 111(a) and 37 C.F.R. 1.53(b) is a new utility patent application for an invention entitled:

SCENE CHANGE MARKING FOR THUMBNAIL EXTRACTION

and invented by:

WILLIAM R. BELKNAP

If a **CONTINUATION APPLICATION**, check appropriate box and supply the requisite information:☐ Continuation ☐ Divisional ☐ Continuation-in-part (CIP) of prior application No.: _____

Which is a:

☐ Continuation ☐ Divisional ☐ Continuation-in-part (CIP) of prior application No.: _____

Which is a:

☐ Continuation ☐ Divisional ☐ Continuation-in-part (CIP) of prior application No.: _____

Enclosed are:

Application Elements

1. ☒ Filing fee as calculated and transmitted as described below
2. ☒ Specification having 22 pages and including the following:
 - a. ☒ Descriptive Title of the Invention
 - b. ☐ Cross References to Related Applications (if applicable)
 - c. ☐ Statement Regarding Federally-sponsored Research/Development (if applicable)
 - d. ☐ Reference to Microfiche Appendix (if applicable)
 - e. ☒ Background of the Invention
 - f. ☒ Brief Summary of the Invention
 - g. ☒ Brief Description of the Drawings (if drawings filed)
 - h. ☒ Detailed Description
 - i. ☒ Claim(s) as Classified Below
 - j. ☒ Abstract of the Disclosure

UTILITY PATENT APPLICATION TRANSMITTAL
(Large Entity)

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Docket No.
ST9-99-180

Total Pages in this Submission

Application Elements (Continued)

3. ☒ Drawing(s) *(when necessary as prescribed by 35 USC 113)*
- a. ☒ Formal Number of Sheets 6
- b. ☐ Informal Number of Sheets _____
4. ☒ Oath or Declaration
- a. ☒ Newly executed *(original or copy)* ☐ Unexecuted
- b. ☐ Copy from a prior application (37 CFR 1.63(d)) *(for continuation/divisional application only)*
- c. ☒ With Power of Attorney ☐ Without Power of Attorney
- d. ☐ DELETION OF INVENTOR(S)
Signed statement attached deleting inventor(s) named in the prior application,
see 37 C.F.R. 1.63(d)(2) and 1.33(b).
5. ☐ Incorporation By Reference *(usable if Box 4b is checked)*
The entire disclosure of the prior application, from which a copy of the oath or declaration is supplied under
Box 4b, is considered as being part of the disclosure of the accompanying application and is hereby
incorporated by reference therein.
6. ☐ Computer Program in Microfiche *(Appendix)*
7. ☐ Nucleotide and/or Amino Acid Sequence Submission *(if applicable, all must be included)*
- a. ☐ Paper Copy
- b. ☐ Computer Readable Copy *(identical to computer copy)*
- c. ☐ Statement Verifying Identical Paper and Computer Readable Copy

Accompanying Application Parts

8. ☒ Assignment Papers *(cover sheet & document(s))*
9. ☐ 37 CFR 3.73(B) Statement *(when there is an assignee)*
10. ☐ English Translation Document *(if applicable)*
11. ☒ Information Disclosure Statement/PTO-1449 ☒ Copies of IDS Citations
12. ☐ Preliminary Amendment
13. ☒ Acknowledgment postcard
14. ☒ Certificate of Mailing
- ☐ First Class ☒ Express Mail *(Specify Label No.):* EL 665 962 503 US

UTILITY PATENT APPLICATION TRANSMITTAL
(Large Entity)

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Docket No.
ST9-99-180

Total Pages in this Submission

Accompanying Application Parts (Continued)

15. ☐ Certified Copy of Priority Document(s) *(if foreign priority is claimed)*

16. ☐ Additional Enclosures *(please identify below):*

Fee Calculation and Transmittal

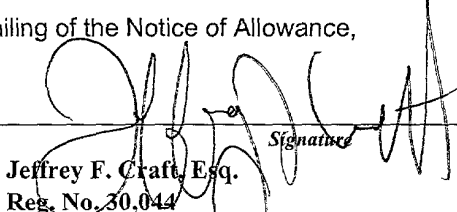
CLAIMS AS FILED

For	#Filed	#Allowed	#Extra	Rate	Fee
Total Claims	45	- 20 =	25	x \$18.00	\$450.00
Indep. Claims	3	- 3 =	0	x \$78.00	\$0.00
Multiple Dependent Claims (check if applicable) <input type="checkbox"/>					\$0.00
BASIC FEE					\$710.00
OTHER FEE (specify purpose)					\$0.00
TOTAL FILING FEE					\$1,160.00

- ☐ A check in the amount of _____ to cover the filing fee is enclosed.
- ☒ The Commissioner is hereby authorized to charge and credit Deposit Account No. 09-0460 as described below. A duplicate copy of this sheet is enclosed.
- ☒ Charge the amount of \$1,160.00 as filing fee.
 - ☒ Credit any overpayment.
 - ☒ Charge any additional filing fees required under 37 C.F.R. 1.16 and 1.17.
 - ☐ Charge the issue fee set in 37 C.F.R. 1.18 at the mailing of the Notice of Allowance, pursuant to 37 C.F.R. 1.311(b).

Dated: November 10, 2000

cc:


Signature
Jeffrey F. Craft, Esq.
Reg. No. 30,044
Pretty & Schroeder, P.C.
444 S. Flower St., 19th Floor
Los Angeles, California 90071
Tel: (213) 622-7700
Fax: (213) 489-4210

UTILITY APPLICATION

OF

William R. Belknap

FOR

UNITED STATES LETTERS PATENT

on

SCENE CHANGE MARKING FOR THUMBNAIL EXTRACTION

Docket No. ST9-99-180 / IBMST 44274

Drawings: 6

Attorneys

PRETTY & SCHROEDER, P.C.

444 South Flower Street, 19th Floor

Los Angeles, California 90071

Ofc: 213/622-7700

Fax: 213/489-4210

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

"EXPRESS MAIL" MAILING LABEL NUMBER EL 665 962 503 US

DATE OF DEPOSIT November 10, 2000

I HEREBY CERTIFY THAT THIS PAPER OR FEE IS BEING DEPOSITED WITH THE UNITED STATES POSTAL SERVICE "EXPRESS MAIL POST OFFICE TO ADDRESSEE" SERVICE UNDER 37 CFR 1-10 ON THE DATE INDICATED ABOVE AND IS ADDRESSED TO BOX PATENT APPLICATION, THE ASSISTANT COMMISSIONER FOR PATENTS, WASHINGTON, D. C. 20231

Joanne H. Galvez

(TYPED OR PRINTED NAME OR PERSON MAILING PAPER OR FEE)



(SIGNATURE OF PERSON MAILING PAPER OR FEE)

SCENE CHANGE MARKING FOR THUMBNAIL EXTRACTION

BACKGROUND OF THE INVENTION

1. Field of the Invention.

This invention relates in general to video transmission systems performed by computers,
5 and in particular, to marking scene changes in video streams.

2. Description of Related Art.

Various applications process video streams or video feeds which are presented to users for
different applications. For example, a newsroom may receive video streams from a satellite or
other transmission device relating to a recent news event or an Internet user may request a video
10 stream or story board of a sports event. These video streams or files must be processed and
presented to the user as quickly as possible since they relate to recent news events and/or must be
transmitted over a network quickly to minimize the waiting time assumed by the user.

Before the video stream or file can be presented to a user, however, the video streams are
processed such that certain frames that represent the most relevant information of the video stream
15 are selected. These frames are often coined "scene change" frames as compared to other frames
which may portray negligible content differences from previous frames. In this context, a scene
change occurs when the content of a first frame of the video stream changes sufficiently in a
second frame of the video stream such that the second frame triggers a new view relative to the
first frame. In order to generate the requested video streams or files, the video streams are
20 processed and analyzed to identify and select scene change frames such that the frames ultimately
presented to the user contain the most relevant information.

Examples of applications using scene change analysis to select frames include newsroom
videos and Internet files. In the context of newsrooms or news editors and producers, video
streams relating to recent news stories may be received from a satellite, a live feed, or a video tape
25 in analog or digital video format. These video streams are analyzed to identify the scene change
frames, and these frames are selected and compiled into, for example, a video clip. As these
streams may relate to recent news items, this processing and selection must be completed as

quickly as possible to insure that the resulting video file is played when the news story is still significant.

Similarly, in the context of the Internet, users may request video files or a storyboard which invokes an extraction tool to select frames of a video file. A storyboard is a collection of images or a collection of thumbnails (i.e., smaller images representing scenes from a video file). An extraction tool, such as a thumbnail extraction tool, may be used to create the storyboard.

More specifically, the Internet is a collection of computer networks that exchanges information via Transmission Control Protocol/Internet Protocol ("TCP/IP"). The Internet computer network consists of many Internet networks, each of which is a single network that uses the TCP/IP protocol. Via its networks, the Internet computer network enables many users in different locations to access information (e.g., video streams) stored in data sources in different locations.

The World Wide Web (i.e., the "WWW" or the "Web") is a hypertext information and communication system used on the Internet computer network with data communications operating according to a client/server model. Typically, a Web client computer will request data stored in data sources from a Web server computer, at which Web server software resides. The Web server software interacts with an interface connected to, for example, a Database Management System ("DBMS"), which is connected to the data sources. These computer programs residing at the Web server computer will retrieve and transmit the data, including video data, to the client computer. Many video streams are transformed into video files that follow digital video compression standards and file formats developed by the Motion Pictures Experts Group (MPEG). These are referred to as MPEG files and are typically files corresponding to movies. Furthermore, there are various video file formats, including MPEG-1, MPEG-2, and MPEG-4 which produce video files at different resolutions.

Some users request storyboards that are comprised of frames of a video file. When a storyboard is to be generated from a video stream or video clip, an application calls a thumbnail extraction tool to conduct scene change analysis and determine which frames of an MPEG file should be selected as part of the storyboard, i.e., which frames were selected from a video stream as scene change frames. Scene change analysis in this context involves comparing a first frame of an MPEG file to a second frame of the MPEG file, etc. for each pair of frames. Frames representing scene changes are selected by the thumbnail extraction tool based on different factors

(e.g., the degree of pan, scan, zoom, etc.) and these selected video stream frames are compiled into a video file. Each of these frames may be an image or "thumbnail" in the storyboard. An MPEG file may include thousands or tens of thousands of frames. Since the storyboard frames must be selected and presented quickly such that the user can select the frames shortly after choosing to generate a storyboard, the analysis and selection of the frames to include within the storyboard must be done as quickly as possible.

Thus, it is clear that video streams must be processed as quickly as possible whether the video stream will ultimately become a news clip, part of an MPEG file or storyboard, or used within some other application or file. In processing video streams or video files, conventional systems process frames twice to determine which frames of the video stream to include within a particular video file. First, when the video stream is initially encoded, frames are processed, for example, to add closed captioning. Second, when an application requests, for example, a storyboard, an extraction tool processes the frames to determine which frames will be selected to create the storyboard. The extra time required to process the frames a second time is assumed by the user. The problem is even more troublesome when there are multiple requests for a video file to create different storyboards. For example, if five different applications request storyboards based on different criteria, a thumbnail extraction tool must perform the scene change analysis five separate times to determine which frames to include within the storyboard for each application.

As illustrated by these simple examples, conventional systems do not access scene change data in real time or near real time, and thus, are inefficient. Consequently, the cost to process video streams is substantially increased. These shortcomings are amplified when scene change analysis is performed manually or by a slower, more complicated system. In addition, if a system is configured to recognize finer changes between scenes, substantially more time may be required to perform the scene change analysis since these more detailed analyses may involve more complicated calculations.

Thus, there is a need in the art providing scene change analysis to extraction tools in real time or near real time for different video stream applications.

SUMMARY OF THE INVENTION

To overcome the limitations in the prior art described above, and to overcome other limitations that will become apparent upon reading and understanding the present specification, the present invention discloses a method, apparatus, and article of manufacture for marking scene change data.

According to an embodiment of the invention, a video stream with multiple frames is received by a computer. The frames of the video stream are analyzed to identify scene changes between the frames. Each frame of the video stream includes a field that can be marked with scene change data. The fields of the frames of the video stream representing scene changes are marked.

BRIEF DESCRIPTION OF THE DRAWINGS

Referring now to the drawings in which like reference numbers represent corresponding parts throughout:

FIG. 1 schematically illustrates an environment in which the scene change marking system may execute;

FIG. 2 schematically illustrates a network environment in which the processed video stream may be used for an embodiment of the present invention;

FIG. 3 is a flow diagram illustrating a process that uses the scene change marking system;

FIG. 4 is a flow diagram illustrating another process that uses the scene change marking system;

FIG. 5 is a flow diagram illustrating an additional process that uses the scene change marking system; and

FIG. 6 is a flow diagram illustrating the use of a processed video file marked by the scene change marking system.

DETAILED DESCRIPTION

In the following description of an embodiment of the invention, reference is made to the accompanying drawings which form a part hereof, and in which is shown by way of illustration a specific embodiment in which the invention may be practiced. It is to be understood that other

embodiments may be utilized and structural and functional changes may be made without departing from the scope of the present invention.

Scene Change Marking For Thumbnail Extraction

Referring to FIG. 1, an illustration of an environment 100 in which one embodiment of the present invention may be used, initially, a video stream or video feed 110 including multiple frames is received from a satellite transmission, a live feed, a video tape, or other source in analog or digital format. The unprocessed, video stream 110 is read into an encoder 120. Alternatively, the video stream 110 may be read into a file or other storage device and then read into an encoder 120. The encoder 120 receives the video stream 110 and analyzes each frame to determine which frame or frames represent a scene change by using an internal 122 or external 124 scene change detection device. The scene change marking system 126, associated with the scene change marking devices 122 and 124, marks one or more fields (e.g., data fields) of one or more frames identified by the scene change detection devices 122 and 124. After the scene change analysis is performed, the encoder 120 outputs either a different video stream or a video file 130. The scene change analysis may be performed within the encoder before, during, or after other processing tasks, for example, compression.

The video file 130 may be compressed into, for example, a Motion Pictures Experts Group (MPEG) format, while retaining the frames with marked fields (e.g., data fields) representing scene changes. Thus, the video file 130 may include full frames 132 and "delta" frames 134. A delta frame 134 includes a portion of a full frame 132. For purposes of illustration, in FIG. 1, a full frame 132 is illustrated as a shaded frame, whereas a delta frame 134 is illustrated as a non-shaded frame. The frames of the video stream 110 may include a field 136. The field 136 may be a user data field. In an alternative embodiment of the present invention, the field 136 may be a private data field. The following description refers to user data fields, although private data fields or other fields could also be utilized. Thus, for one embodiment of the invention, each full frame 132 and each delta frame 134 includes user data fields 136.

Typically, there are multiple user data fields 136, and these user data fields 136 are used for various purposes. For example, conventional systems use the user data field to store closed captioning information. In addition to this data, the scene change marking system 126 can also utilize the user data field 136 to update frames which represent scene changes with scene change

data. For purposes of illustration, an updated user data field 136 is indicated with a dark box in the upper left corner of each frame. The video file 130, which may include scene change frames with marked user data fields 136 may be stored in data store 140. An extraction tool invoked by an application may then access the processed video file 130 stored in data store 140 to create a storyboard. In particular, the extraction tool selects frames from the processed video file 130 by referring to the scene change data marked in the user data field 136 of each frame.

Those skilled in the art will recognize that the exemplary environment illustrated in Figure 1 is not intended to limit the present invention. Indeed, those skilled in the art will recognized that other alternative system environments may be used without departing from the scope of the present invention. For example, in one embodiment, the scene change marking system 126 is part of an encoder 120. However, other alternative embodiments may be configured such that the scene change marking system 126 is independent of the encoder but still associated with the encoder or such that the scene change marking system includes all of the components of Figure 1 and coordinates the activities of each component.

FIG. 2 schematically illustrates a network environment 200 in which a video file 130 processed by one embodiment of the present invention. In particular, Figure 2 illustrates a typical distributed computer system 200 using a network 202 to connect client computers 204 executing client applications to a server computer 206 executing software and other computer programs, and to connect the server system 206 to data sources 208. A data source 208 may comprise, for example, a multi-media database storing video files in MPEG format with marked fields (e.g., a processed video file 130). These processed video file 130 may alternatively be stored at one or more of the client computers 204.

A typical combination of resources may include client computers 204 that are personal computers or workstations, and a server computer 206 that is a personal computer, workstation, minicomputer, or mainframe. These systems are coupled to one another by various networks, including LANs, WANs, SNA networks, and the Internet. Each client computer 204 and the server computer 206 additionally comprise an operating system and one or more computer programs. The server computer 206 also uses a data source interface and possibly, other computer programs, for connecting to the data sources 208. The client computer 204 is bi-directionally coupled with the server computer 206 over a line or via a wireless system. In turn, the server computer 206 is bi-directionally coupled with data sources 208.

An extraction tool may be a software program resident on a client computer 204 or a server computer 206 connected to a network 202. Those skilled in the art will recognize that the extraction tool may also be implemented on hardware or firmware. A client computer 204 typically executes a client application (e.g., a browser) and is coupled to a server computer 206
5 executing one or more server software programs. The extraction tool may then be directly invoked by a user or invoked by another application at the client computer 204. If the extraction tool were resident at the client computer 204, the extraction tool would access a processed video file 130 stored in a data source 208 via the server computer 206. Then, the extraction tool would use the marked fields 136 to extract frames representing scene changes for generating thumbnails.

10 In another embodiment, the server software may include the extraction tool. In this case, the extraction tool would access processed video file 130 stored in a data source 208 and create a storyboard. If the extraction tool were invoked by a user or application at the client computer 204, the extraction tool would transmit the storyboard to the client computer 204. Additionally, the server software may use the scene change data to generate an index of access points for
15 displaying specific scenes or segments.

The operating system and computer programs are comprised of instructions which, when read and executed by the client and server computers 204 and 206, cause the client and server computers 204 and 206 to perform the steps necessary to implement and/or use the scene change marking system 126. Generally, the operating system and computer programs are tangibly
20 embodied in and/or readable from a device, carrier, or media, such as memory, other data storage devices, and/or data communications devices. Under control of the operating system, the computer programs may be loaded from memory, other data storage devices and/or data communications devices into the memory of the computer for use during actual operations.

25 Thus, the present invention may be implemented as a method, apparatus, or article of manufacture using standard programming and/or engineering techniques to produce software, firmware, hardware, or any combination thereof. The term "article of manufacture" (or alternatively, "computer program product") as used herein is intended to encompass a computer program accessible from any computer-readable device, carrier, or media. Of course, those skilled in the art will recognize many modifications may be made to this configuration without departing
30 from the scope of the present invention.

Those skilled in the art will also recognize that the exemplary environment 200 illustrated in FIG. 2 is not intended to limit the present invention. Indeed, those skilled in the art will recognize that other alternative hardware environments may be used without departing from the scope of the present invention. The manner in which the present scene change marking system 126 operates within both of the previously described environments 100 and 200 will be explained below in further detail.

FIG. 3 is a flow diagram illustrating a scene change marking technique as implemented by the scene change marking system 126. In block 300, the unprocessed, video stream 110 is received by an encoder 120. If the unprocessed video stream 110 is in analog format, the encoder 120 converts the analog video stream 110 into a digital video stream.

In block 310, the encoder 120, with an internal 122 or external 124 scene change detection device, analyzes each frame of the motion-based video 110. The encoder 120 compares a first frame to a second frame, the second frame to a third frame, etc. for every frame within the video stream 110 to identify which frames represent scene changes. A scene change occurs when the content of a frame changes sufficiently compared to a previous frame to trigger a new view. A scene change normally occurs between two full frames 132. In alternative embodiments, a scene change may occur between a full frame 132 and a delta frame 134 or between two delta frames 134. Scenes that represent scene changes typically portray significant changes in the content of the video stream 110 whereas "non scene change" frames portray less significant details (e.g., details that are repeated from prior frames). These details may be extracted from previous frames and therefore do not need to be encoded, and will be filled in or smoothed out by the human eye. Thus, the encoder 120, with the scene change marking system 126, identifies scene changes in frames to determine which frames are "main" or "key" frames.

A scene change may be triggered by any number of changes within a scene. For example, a scene change may be based on a degree of change caused by a pan or a scan, a tilt, a zoom, a cut, and other changes. A pan or a scan involves moving a camera along a horizontal axis to follow action and to reveal the content of a scene to the audience. In other words, strips of a scene are deleted from one or both sides of a picture and a corresponding number of strips are added to the frame. Thus, when panning is done across a scene that is larger than the screen, most of the image does not change. A tilt, however, refers to moving a camera in a vertical motion. A zoom may "zoom in" from or "zoom out" from an object. When the magnification of the objects are

increased, the viewer "zooms in" on the object. If the magnification of the objects decrease, the viewer "zooms out" from the object. In a zoom, the relative positions and sizes of all of the objects remain the same. A cut involves a change in either camera angle, placement, location, or time. In other words, a cut is an abrupt change into a new scene.

Further, scene change frames may be selected based on application specific criteria or criteria selected by a user. For example, an application or user may specifically need frames that represent scene changes based on a zoom of 150%. In this example, the scene change analysis will focus on zooms rather than other scene change attributes such as pans, scans or cuts. Thus, applications or users may specify any number and combination of scene change attributes to represent scene changes.

In addition, scene change analysis may be performed by automated devices. Automated devices may utilize techniques that perform complex calculations to analyze scene changes. A typical video stream 110 may include thousands or tens of thousands of frames. Thus, analyzing every frame of the video stream 110 can easily involve numerous complicated calculations. The calculations become even more complex if the internal or external scene detection device 122 or 124 is tuned to read and analyze finer changes within a scene. For example, computations may be more complex if a scene change is defined when an object moves an inch as compared to a foot or when an object is zoomed by 100% as compared to 200%.

Continuing with FIG. 3, in block 320, the user data fields 136 of scene change frames are updated by the scene change marking system 126. With reference to FIG. 1, an updated user data field 136 is represented as a dark box in the upper left corner of a frame (i.e., frames 1 and 4).

The user data field 136 may store various types of scene change data. For example, conventional systems utilize the user data field 136 to store closed captioning data. The amount of memory allocated to each user data field 136 depends on the complexity of the technique that identifies scene changes. Less space may be available for the user data field 136 if the scene change analysis involves more complex analysis.

In contrast to conventional systems, the scene change marking system 126, utilizes the user data field 136 to store scene change data. As previously explained, in an alternative embodiment, private data fields may also be used to store updated scene change data. Thus, when the encoder 120 identifies scene changes within the motion-based video 110 during its initial analysis of the video stream 110, the scene change marking system 126 updates the field 136 of each frame

representing a scene change with appropriate scene change data concurrently with this initial analysis.

The marking of the frames to indicate a scene change must be done in a transparent manner for the encoded content. For example, in MPEG, optional “user data” can be held in each encoded video frame. The decoders will either utilize the data if they are programmed to do so, or they will discard it without impact to the rest of the decoding process. The scene change marking system implements this technique with a signal or user data start code of hex: 00 00 01 B2. The start code cannot contain 23 consecutive zeros. The first user data start code of hex is terminated by the next start code. One possible implementation of a user data field would be to have a word indicator signaling Scene_Change_Indicator and a word field indicating the type of scene change. For example, the type of scene change may be indicated such that 01 = new, 02 = pan, 03 = scan., etc. A more complex example would add a percentage field that indicates the percentage of change in the scene, a direction field for overall motion changes, or an effects indicator to describe a standard transition effect.

The flexibility of the user data field configurations to signify scene changes presents a number of alternative embodiments for the structure of the user data field 136. For example, in an alternative embodiment, the user data field 136 is structured such that a single data bit indicates whether a scene change occurred, and thus, whether the corresponding frame should ultimately be selected.

In yet another alternative embodiment of the present invention, the data bits may indicate that a scene change occurred due to one or more specific scene change attributes. For example, four data bits may represent four different scene change attributes (e.g., bit 1 may represent a scan, bit 2 may represent a tilt, bit 3 may represent a zoom, and bit 4 may represent a cut). If a camera is tilted sufficiently such that a scene change occurs in a frame, bit 2 of the user data field 136 of that frame, corresponding to a tilt, may be updated to indicate that the frame represents a scene change. However, if the scene change was caused by a cut, bit 4 may be updated. Further, one skilled in the art will recognize that any number of data bits may identify scene changes caused by changes of specific scene change attributes.

In an additional alternative embodiment of the present invention, one or more additional data bits may be allocated within the user data field 136 to represent a quantify or amount of change caused by a corresponding scene change attribute. For example, additional data bits may

indicate that a 25 degree tilt or a 150% zoom occurred within a frame. The encoder 120 or other scene detection device may then interpret these quantities and determine whether a scene change occurred. The corresponding data bits within the user data field 136 of that frame may be updated to indicate that a scene change occurred within that frame.

Those skilled in the art will recognize that the scope of the invention encompasses a user data field 136 that may utilize different data configurations to represent scene change data or to indicate that scene change occurred. Therefore, the previously described embodiments are not intended to limit the present invention.

Further, in additional embodiments of the present invention, within the encoder 120, the scene change marking system 126 may mark user data fields 136 of frames representing scene changes during or after the video file 130 is compressed as explained below.

Referring to FIG. 3, in block 330, the video file 130 is compressed into, for example, an MPEG file 130. The MPEG file 130 includes a fraction of the total number of full frames within the video stream 110. To illustrate scene changes with an example, a video stream 110 of a football game may include a full frame 132 with an image illustrating a person ready to throw the ball and the stadium background. The video stream 110 may also include full frames 132 representing changes from the initial full frame 132 illustrating the starting pose of throwing a ball. Each subsequent frame includes images of the quarterback's hand in slightly different positions as compared to the previous full frame 132 to emulate the arm moving forward to throw the ball. If the cameras are moved or switched to a different video of the quarterback and stadium, a scene change has occurred, and this change is signified in the user data field 136 of the corresponding frame.

As part of the compression process, the frames representing scene changes may be retained in the MPEG file 130 whereas other frames are deleted or not encoded. Thus, if the stadium background does not change, the background does not be illustrated in the delta frames 134 which represent only the changes between frames. In this scenario, the delta frames will not be encoded as part of the compression process performed by the encoder 120. However, as the frames progress and portray the person's arm moving forward further, a delta frame 134 may illustrate an image of the arm in different positions as the forward motion is completed such that a scene change has occurred. This frame may be included within the MPEG file 130. If this process is continued for the remaining frames portraying the quarterback throwing the ball, the series of full

frames 132 and delta frames 134 displayed in quick succession creates an illusion of a quarterback throwing a football.

Since not all details can be captured and stored, there may be missing information in the video file. Because only some "important" details are selected, there are "gaps of detail" in the video file. In other words, the compressed frames portray the image of a quarterback throwing the ball and any discontinuities resulting from lost image content may be "smoothed out" or "filled in" by the human eye which still perceives an illusion of throwing a ball when the frames are viewed in succession. Smaller images of scene change frames, often termed "thumbnails," may also be generated to represent different portions of the video file 130 to create a storyboard.

The resulting MPEG file 130 may be compressed to as little as 1% of the original video stream file size. Eliminating this much data may be necessary because of limited transmission bandwidth, some of which is also consumed by related audio and sound files. Further, the smaller size file of video file 130 is desired to reduce the time required to download the file. This problem is amplified if slower modems or network connections are utilized. The resulting compressed video file 130 includes multiple frames in the same format (e.g., MPEG or Joint Photographic Experts Group (JPEG) format) which store changes from one frame to another instead of storing each entire frame of the original motion-based video 110. The frames storing changes between frames are delta frames 134 which are based on prior full frames 132. After the video stream is processed and a compressed video file 130 is generated, the video file, including the frames with marked user data fields 136 are stored to a data store 140 in block 340.

FIG. 4 is a flow diagram illustrating another process that uses the scene change marking system 126. In an alternative embodiment of the present invention, instead of using the technique illustrated in FIG. 3, the scene change marking system 126 can be utilized by performing block 400 in which an encoder 120 receives an unprocessed, video stream 110. Then, in block 410, the video stream 110 frames are analyzed to identify scene changes. In block 420, the video stream 110 is compressed. Continuing with block 430, user data fields 136 of frames representing scene changes are marked. Finally, in block 440, the video file 130 with the marked fields is stored in a data store 140.

FIG. 5 is a flow diagram illustrating an additional process that uses the scene change marking system. In yet another alternative embodiment of the present invention, initially, in block 500, an unprocessed video stream 110 is received by an encoder 120. In block 510, the video

stream 110 is compressed. Then, in block 520, the video stream 110 is analyzed to identify scene changes. In block 530, the fields 136 representing scene changes are marked. Finally, in block 540, the video file 130 containing the marked fields is stored in a data store 140.

Considering the possible techniques illustrated in FIGS. 3, 4, and 5, scene change analysis can be performed in either the compressed domain or the uncompressed domain while the field marking may occur in the compressed domain.

FIG. 6, a flow diagram illustrating one application of the present invention to generate a storyboard. In block 600, an extraction tool receives a request for thumbnails. The thumbnails must be retrieved from the marked video file 130 that is stored in a data store 140. The video file 130 includes both scene change frames and other frames. All of the frames may include user data fields 136, however, typically only some of the frames have user data fields 136 that were marked by the scene change marking system 126. The extraction tool is configured to read scene change data in the user data fields 136 of all of the frames of the video file 130 to determine which frames to select instead of repeating the scene change analysis as required in conventional systems.

By reading the user data fields 136 of each frame, i.e., reading the scene change data in real time or near real time, the extraction tool determines which user data fields 136 were marked, and thus, which frames represent scene changes. In block 610, the extraction tool extracts scene change frames from the video file 130. Thus, with the scene change marking system 126, the extraction tool is not required to repeat the scene change analysis to identify scene change frames. After the extraction tool selects the scene change frames, in block 620, the extraction tool transfers the scene change frames to the application.

Considering the forgoing description, the scene change marking system 126 overcomes the limitations in conventional systems caused by the inability to access scene change data in real-time or near real-time. First, the scene change marking system 126 eliminates or minimizes the time required to identify scene change frames by enabling an extraction tool or other application to read the user data fields 136 to identify scene change frames. As a result, it is not necessary to repeat the complicated scene change analysis.

The time saved by the scene change marking system 126 can be significant. For example, conventional extraction tools may require 10 minutes to process and analyze scene changes within a 40-50 minute video clip. By accessing scene change data in the user data fields 136 in real time,

the scene change marking system 126 eliminates this delay which would otherwise be assumed by the user who requested the file 130.

Further, the time saved by not repeating the scene change analysis also benefits newsroom editors and producers who receive video streams 110 and are required to generate video files 130 of recent news events in a short period of time. These video files 130 may include numerous scene changes, and thus, scene change analysis may be a bottleneck to processing the video. The scene change marking system 126 eliminates this bottleneck by saving editors and producers critical processing and editing time that is otherwise lost by repeating the intensive scene change analysis. Rather, applications can be configured to read the user data fields 136 to identify scene change frames. In addition, since applications involving the Internet may process and download video files more slowly, more users may choose to use the Internet to download storyboard files since downloads may be completed in a shorter period of time. Therefore, an extraction tool with real time access to scene change data can enhance numerous different applications involving video feeds, whether those applications relate to news publications, video editing, Internet, or other applications.

A further advantage of the scene change marking system 126 over conventional systems is that the scene change marking system 126 reduces costs of generating video files 130. Producers and editors generate video clips more efficiently thereby reducing the processing and production costs since the scene change data can be accessed in real time or near real time. Editing tasks that may have previously required hours may be completed within a matter of minutes, and thus, editing and processing costs are significantly reduced.

Moreover, the scene change marking system 126 is advantageous since the quality and accuracy of the resulting video file 130 will be enhanced. Conventional systems analyze the original video feed at the encoder, and then repeat the analysis on the video feed or video files which output by the encoder, i.e., on data that was previously processed. As a result, the resulting video file may omit original movie data. The scene change marking system 126 overcomes this shortcoming by analyzing scene changes only one time rather than repeating the scene change analysis on data that was previously processed. As a result, the scene change marking system 126 generates a more accurate, higher quality video file 130 with improved resolution.

[illegible]

5

10

WHAT IS CLAIMED IS:

1 1. A method of processing a video stream received by a computer, the method
2 comprising:
3 receiving a video stream, wherein the video stream comprises multiple frames;
4 analyzing the video stream to identify scene changes between frames of the video stream;
5 and
6 marking one or more fields of one or more frames of the video stream to indicate a scene
7 change.

1 2. The method of claim 1, wherein the computer comprises an encoder.

1 3. The method of claim 2, wherein marking one or more fields occurs within the
2 encoder.

1 4. The method of claim 1, wherein the field comprises a user data field.

1 5. The method of claim 1, wherein the field comprises a private data field.

1 6. The method of claim 1, wherein a scene change occurs when the content of a first
2 frame of the video stream changes sufficiently in a second frame of the video stream such that the
3 second frame triggers a new view relative to the first frame.

1 7. The method of claim 1, wherein the scene change data comprises a data bit, and
2 wherein updating the data bit indicates a scene change.

1 8. The method of claim 1, wherein the scene change data comprises one or more data
2 bits representing a scene change attribute, and wherein updating one or more data bits indicates
3 a scene change due to a corresponding scene change attribute.

1 9. The method of claim 8, wherein the scene change data further comprises one or
2 more additional data bits, wherein the additional data bits indicate the amount a scene has changed
3 in relation to the corresponding scene change attribute.

1 10. The method of claim 1, further comprising compressing the video stream to
2 generate a video file.

1 11. The method of claim 10, wherein a frame of the video file representing a scene
2 change comprises a full frame.

1 12. The method of claim 10, wherein a frame of the video file representing a scene
2 change comprises a delta frame.

1 13. The method of claim 10, further comprising extracting one of more frames
2 representing a scene change from the video file with an extraction tool, wherein the extraction tool
3 selects frames representing scene changes by reading scene change data in the fields.

1 14. The method of claim 13, wherein the extraction tool accesses the scene change data
2 in the fields in real time.

1 15. The method of claim 13, further comprising generating a storyboard with the
2 extracted frames.

1 16. An apparatus for processing a video stream, comprising:
2 a computer; and
3 one or more computer programs, performed by the computer, for receiving a video stream,
4 wherein the video stream comprises multiple frames; analyzing the video stream to identify scene
5 changes between frames of the video stream; and marking one or more fields of one or more
6 frames of the video stream to indicate a scene change.

- 1 17. The apparatus of claim 16, wherein the computer comprises an encoder.
- 1 18. The apparatus of claim 17, wherein marking one or more fields occurs within the
2 encoder.
- 1 19. The apparatus of claim 16, wherein the field comprises a user data field.
- 1 20. The apparatus of claim 16, wherein the field comprises a private data field.
- 1 21. The apparatus of claim 16, wherein a scene change occurs when the content of a
2 first frame of the video stream changes sufficiently in a second frame of the video stream such that
3 the second frame triggers a new view relative to the first frame.
- 1 22. The apparatus of claim 16, wherein the scene change data comprises a data bit, and
2 wherein updating the data bit indicates a scene change.
- 1 23. The apparatus of claim 16, wherein the scene change data comprises one or more
2 data bits representing a scene change attribute, and wherein updating one or more data bits
3 indicates a scene change due to a corresponding scene change attribute.
- 1 24. The apparatus of claim 23, wherein the scene change data further comprises one
2 or more additional data bits, wherein the additional data bits indicate the amount a scene has
3 changed in relation to the corresponding scene change attribute.
- 1 25. The apparatus of claim 16, further comprising compressing the video stream to
2 generate a video file.
- 1 26. The apparatus of claim 25, wherein a frame of the video file representing a scene
2 change comprises a full frame.

1 36. The article of manufacture of claim 31, wherein a scene change occurs when the
2 content of a first frame of the video stream changes sufficiently in a second frame of the video
3 stream such that the second frame triggers a new view relative to the first frame.

1 37. The article of manufacture of claim 31, wherein the scene change data comprises
2 a data bit, and wherein updating the data bit indicates a scene change.

38. The article of manufacture of claim 31, wherein the scene change data comprises one or more data bits representing a scene change attribute, and wherein updating one or more data bits indicates a scene change due to a corresponding scene change attribute.

39. The article of manufacture of claim 38, wherein the scene change data further comprises one or more additional data bits, wherein the additional data bits indicate the amount a scene has changed in relation to the corresponding scene change attribute.

1 40. The article of manufacture of claim 31, further comprising compressing the video
2 stream to generate a video file.

1 41. The article of manufacture of claim 41, wherein a frame of the video file
2 representing a scene change comprises a full frame.

1 42. The article of manufacture of claim 41, wherein a frame of the video file
2 representing a scene change comprises a delta frame.

43. The article of manufacture of claim 41 , further comprising extracting one of more frames representing a scene change from the video file with an extraction tool, wherein the extraction tool selects frames representing scene changes by reading scene change data in the fields.

1 44. The article of manufacture of claim 44, wherein the extraction tool accesses the
2 scene change data in the fields in real time.

1 45. The article of manufacture of claim 44, further comprising generating a storyboard
2 with the extracted frames.

ODM F T 36050260

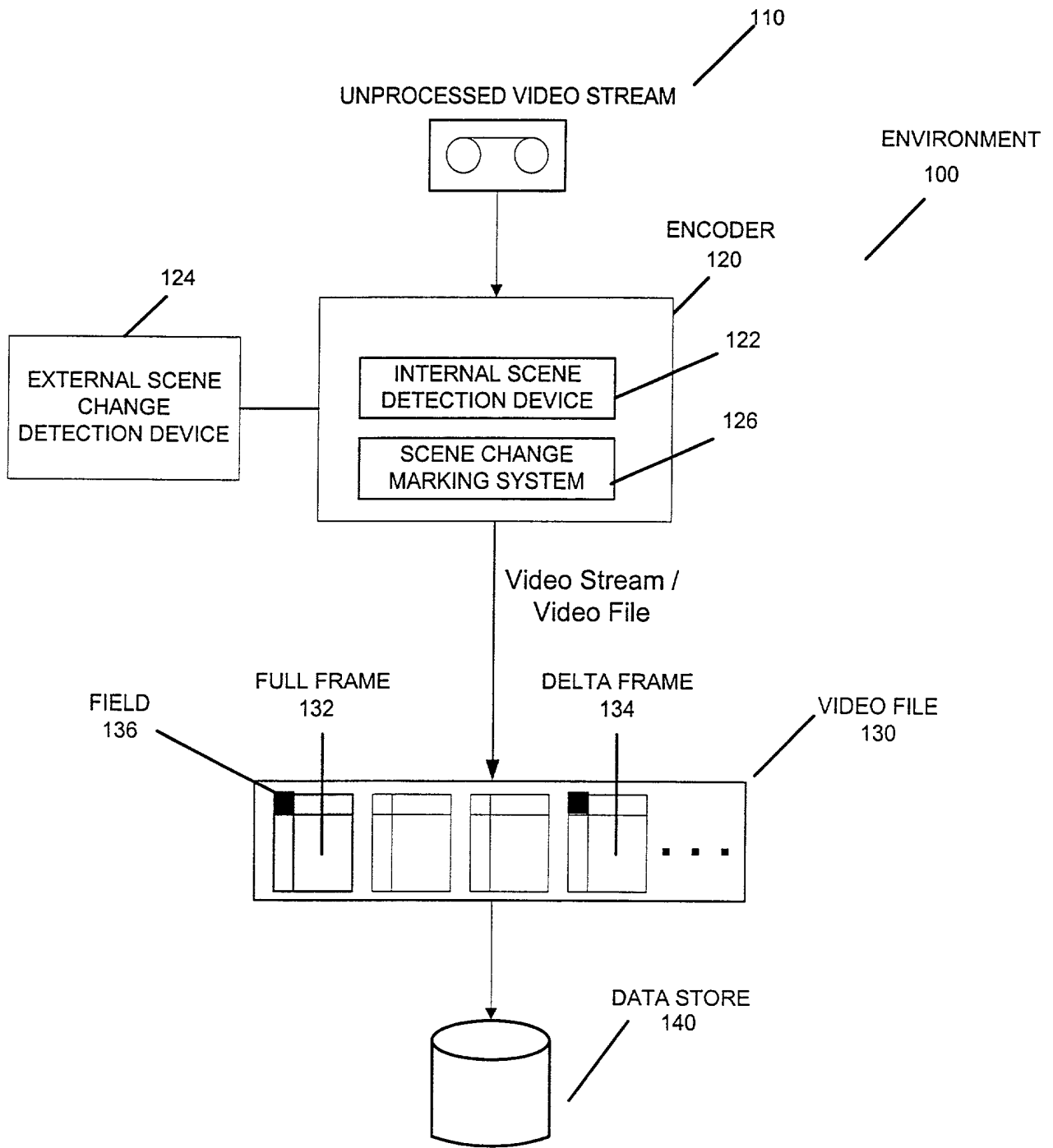


FIG. 1

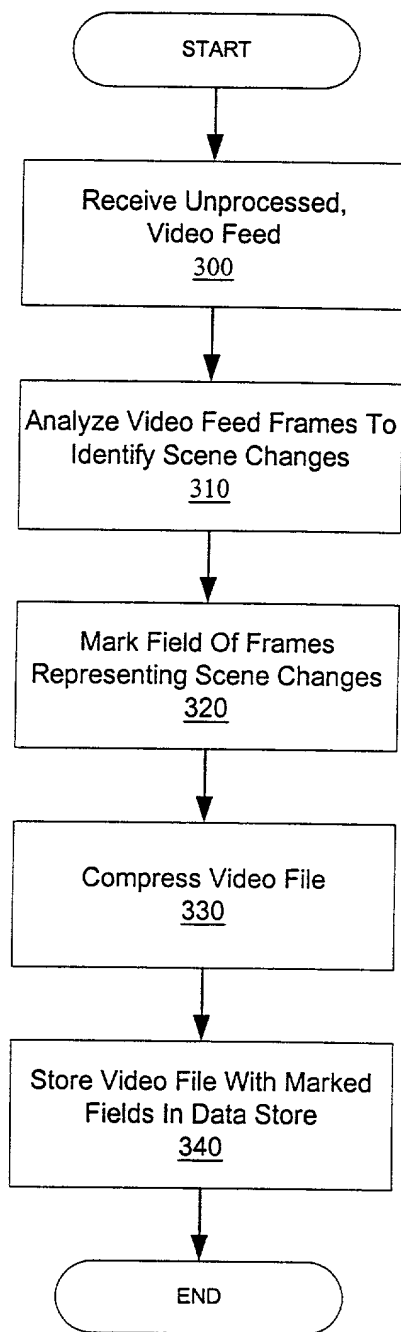
[illegible]

FIG. 3

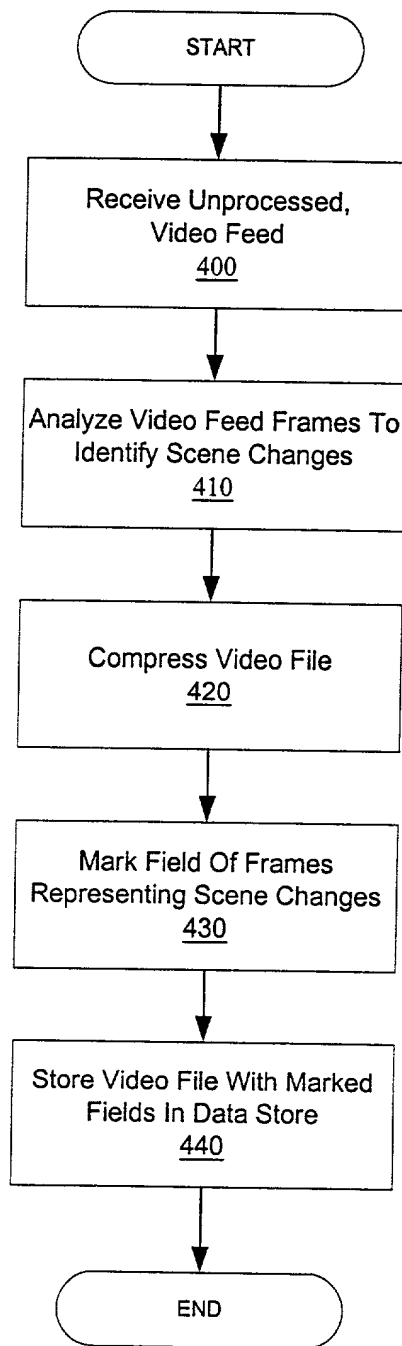


FIG. 4

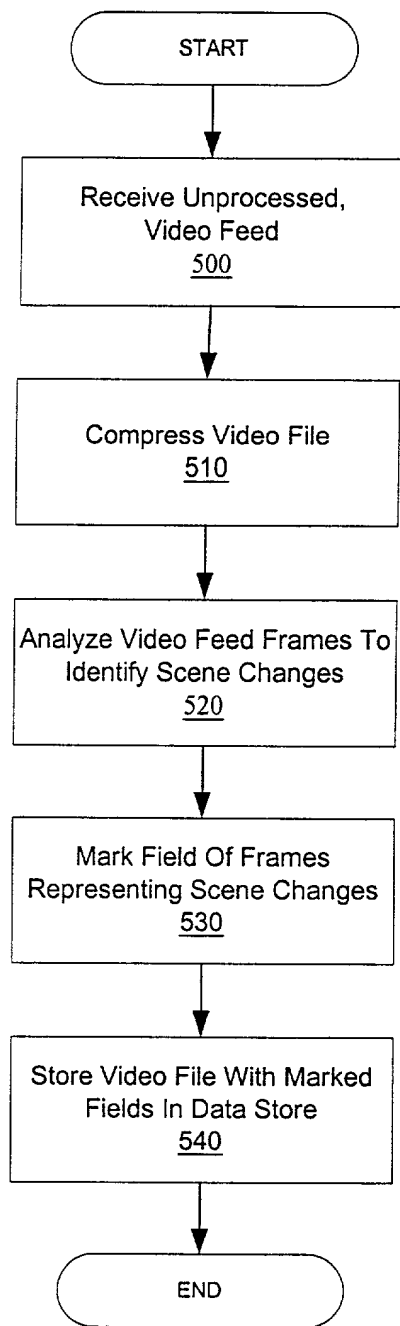


FIG. 5

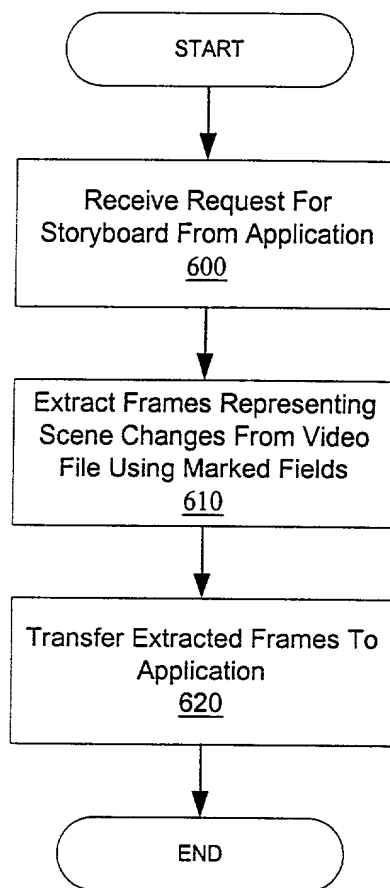


FIG. 6

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

POWER OF ATTORNEY: As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith. (list name and registration number)

Romualdas Strimaitis, Registration No. 35,697, Prentiss W. Johnson, Registration No. 33,123, Ingrid M. Foerster, Registration No. 36,511, Timothy M. Farrell, Registration No. 37,321, Christopher A. Hughes, Registration No. 26,914, John E. Hoel, Registration No. 26,279, Edward A. Pennington, Registration No. 32,588, and Joseph C. Redmond, Jr., Registration No. 18,753, all of whom are attorneys with IBM Corporation; and

Laurence H. Pretty, Registration No. 25,312, Robert A. Schroeder, Registration No. 25,373, Jeffrey F. Craft, Registration No. 30,044, and Janaki Komanduri, Registration No. 40,684, all of whom are attorneys with Pretty & Schroeder.

Send correspondence to: Janaki Komanduri, Esq.
Pretty & Schroeder, P.C.
444 S. Flower Street, 19th Floor
Los Angeles, California 90071

Direct Telephone Calls to: (name and telephone number) Janaki Komanduri, (213) 622-7700

Full name of sole or first joint-inventor: William R. Belknap

Inventor's signature:

Date:

William R. Belknap

10/19/00

Residence: 1262 Blewett Avenue, San Jose, California 95125

Citizenship: United States

Post Office Address: Same